The University of Alberta
Printing Department
Edmonton, Alberta

THE UNIVERSITY OF ALBERTA

CORRELATES OF GRADE ONE ACHIEVEMENT

by

© JAMES WILLIAM IRVINE

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA
OCTOBER, 1968

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Correlates of Grade One Achievement," submitted by James William Irvine, in partial fulfilment of the requirements for the degree of Master of Education.

ABSTRACT

This study was designed to examine the relationships between screening tests and measures of grade one academic achievement. In an attempt to sample skills relevant to formal schooling, a composite battery of published tests, correlates of learning difficulties and intuitively developed simple tests was administered to 42 grade one Edmonton children. From the battery of 34 independent variables, 16 were retained and administered to another sample of 55 children.

Results indicated that teacher rating was the best single predictor of achievement test performance. Frostig's Developmental Test of Visual Perception, Slingerland's Screening Tests for Identifying Children with Specific Language Disability, Dunn's Peabody Picture Vocabulary Test, Wepman's Auditory Discrimination Test, the Wechsler Intelligence Scale for Children Digit Span subtest, age, writing name, drawing a man, and number of words were all significantly related to achievement measures ($p < .05$). The Perceptual Forms Test and the arm extension test were not significantly related to achievement, while the First Grade Screening Test was only a mildly useful predictor.

Multiple correlation coefficients were computed for each achievement measure, using data from the total sample of 97 children. Optimal combinations of predictors, after correction for attenuation,

accounted for from 41-49% of the variance in achievement test performances ($p < .01$).

Implications were discussed for: (a) follow-up studies, (b) individualized and remedial teaching programs, (c) diagnostic and predictive assessments, and (d) the development of educational testing programs.

ACKNOWLEDGMENTS

Several persons and institutions have assisted the writer in this project.

To my chairman and advisor, Dr. D.R. Cameron, I wish to express my sincere thanks, not only for making graduate studies in Canada a rewarding reality, but also for his informed guidance and interest throughout this investigation. Sincere thanks are expressed to Dr. J.K. Bishop and Dr. R.B. Armstrong for their helpful comments and thoughtful criticisms; to Mr. W. Muir for assistance with data analysis; to the principals, teachers and pupils of Coronation, Glenora and Talmud Torah schools, and in particular to Mrs. L. Bishop, for assistance and cooperation in field testing; to my parents, Mr. and Mrs. A.F. Irvine for many hours of valuable clerical assistance; to my wife, Sue, for her continual understanding and encouragement; and to Dr. R. Elliott for encouragement to undertake the project.

Thanks are expressed also to the University of Alberta and to the Government of the Province of Alberta for financial assistance.

The contribution of each is acknowledged with sincere gratitude.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION TO THE PROBLEM

Recent literature in special education reveals an increase in publications pertaining to children with specific learning problems. The expansion of diagnostic services in many areas has identified large numbers of children who apparently are unable to cope with what is expected of them in school. Many of these children have no readily discernible disabilities and are difficult to accommodate in existing regular or special class programs.

Reflecting the growing concern of educators for children with specific learning problems, Bateman (1966) has outlined some of the problems confronting the reader of relevant literature in this area. Her review includes a stimulating section dealing with issues and needed research (pp. 113-114).

Considering the early identification of potentially handicapped learners, Bateman has commented:

> As public school programs for children with learning disorders increase in number, so will the need for screening tests. Several efforts are under way to construct test batteries that are both diagnostic and predictive of future school performance. One possible outcome of such tests would be the screening of all kindergarten or nursery school children and subsequent placement of them in a program designed to prevent or overcome academic disabilities. Very likely factors can be found which correlate highly with learning disorders seen in a clinic population and which are practical for inclusion in a group screening battery (1966, p. 114).

Assuming that early identification and appropriate interventions may minimize the development of secondary problems associated with reactions to failure, it appears that some form of screening school entrants is indicated. If this initial screening can be done by teachers, it would be useful to refer potentially handicapped students for more detailed diagnosis, thereby making more timely use of available diagnostic and consultant personnel.

In considering the types of tasks which may be appropriate for preschool and kindergarten children, the test constructor must assume limited interest span and limited facility with pencil-paper skills on the part of most examinees. Traditional readiness tests (Bryant, 1959; Cattell, 1949; Gates, 1942; Lee & Clark, 1962) have tended to be group-administered, pencil-paper batteries requiring expendable answer booklets. Although Bateman's review suggests the need to develop instruments more imaginative and comprehensive than currently available readiness tests, it is apparent that Bateman advocates group screening. The value of a screening battery for four, five and six year old children would seem limited if screening devices are omitted because they may not be assessed in a group situation.

In recent literature, several instruments have been discussed which may be of value for inclusion in programs for screening school entrants. Very little evidence has been offered to date to indicate the value of some of these devices for predicting school achievement or for providing guidelines for remedial programs. Additionally, the

relationships between various screening measures and academic achievement criteria have not been clearly established. There is undoubtedly a need to critically examine some of the more recent screening programs, and in addition, to develop and examine other simple procedures for detecting potentially handicapped learners.

Accordingly, the purposes of the present study were:

1. To obtain reliability and validity data for a selection of screening devices, using a sample of grade one Edmonton children;
2. To examine, for a sample of grade one Edmonton children, the relationships between optimal combinations of screening devices and achievement test performances, as an initial step towards the development of a screening procedure for school entrants.

CHAPTER II

THEORETICAL ORIENTATION

The developmental theory of Jean Piaget is sufficiently extensive to form the basis of a theoretical framework within which task-specific failures to learn can be considered.

Piaget views intellectual development as the product of an individual's continual striving for mastery over the demands of his dynamic environment. The essence of this position has been succinctly stated by Flavell (1963) who suggests that:

> ...it does make sense to envisage development as a gradual, step-by-step process of structural accrual and change, each structural form necessarily building on its predecessor, yet - by virtue of new increments of assimilation-accommodation activity - going a little beyond it (p. 407).

For Piaget, it is unnecessary to seek motivational mechanisms such as basic need tensions to explain a child's cognitive activity in a given situation. An individual's constant striving for equilibrium is, in itself, a powerful motivating force. Such a position implies that the likelihood of meaningful learning in a given situation will depend, in large measure, upon the appropriateness of the task for each individual who brings to the situation an unique background of experience. Inasmuch as each individual's background is unique, each can be expected to perceive a task differently. The hypothetical relationship between experience and intellectual performance has been discussed extensively by Hunt (1961). Reflecting

the influence of Piaget, Hunt states that:

> ...it is the appropriateness of the match between the circumstances that the child encounters as he develops and the nature of his own intellectual organizations at the time of the encounters that appear to determine in very large part his rate of intellectual development. ...Both (Piaget and Hebb) see early experience as of probably crucial importance in determining both the rate and the final level of ability (p. 357).

In one sense, experience can be viewed as a "history of reinforcement" by which behavioral repertoires are established (Bijou, 1963). Behavior which has in the past been reinforced in a given situation is more likely to be repeated than behavior which has not been reinforced. Bijou thus conceptualizes failure as attributable to limited behavioral repertoires which are in turn the products of intermittent, inadequate or traumatic reinforcement and extinction histories. Following this approach, a diagnostician would tend to focus on detailed, extensive descriptions of current functioning, together with the specific environmental and intrapersonal factors apparently supporting both the adaptive and maladaptive features of observed performance. Learning problems thus analyzed could not be meaningfully attributed to the operation of hypothetical constructs such as "sub-average intelligence," "brain-damage" etc.

Although Bijou (1963), Hunt (1961) and Piaget (1950) stress the crucial role of experience as a factor determining performance in a given situation, it would be unwise to infer that an individual's observed performance in a given situation is dependent exclusively on

his experiential-background depth, in a quantitative sense. Success or failure in a given task is a function of something more than experiential "readiness." Non-intellective and organic factors may contribute to fluctuations and apparent incongruities in the performance of a given individual.

It is important, therefore, to distinguish between failures due to task inappropriateness, from those apparently due to intra-organismic factors such as sub-optimal or supra-optimal arousal, abnormal situational or physiological conditions, or task-avoidance behavior. The development of new cognitive structures, in the Piagetian sense, is unlikely in situations where cognitive effort is expended in avoiding confrontation with a specific situation. With respect to this analysis of task-specific "failure," Bruner commented:

> There is a sharp distinction that must be made between behavior that copes with the requirements of a problem and behavior that is designed to defend against entry into the problem (1966, pp. 3-4).

From the preceding discussion, a useful theoretical framework for considering academic difficulties in primary school children can be developed.

Academic failure can be conceptualized as the outcome of an inappropriate match between the requirements of the task in question, and a student's experiential repertoire and cognitive energy available at that particular time. When a task is inappropriate, a child may expend considerable energy in unsuccessful attempts to relate the

perceived demands of the task to his existing structures. Alternatively, he may avoid the task it if appears to be unchallenging, threatening or irrelevant. Regrettably, the longer a child makes unsatisfactory academic progress, the more difficult it becomes to distinguish failures due to task inappropriateness from failures due to task avoidance.

Each child entering grade one may be assumed to possess an unique repertoire of skills and habits and an unique set of expectations. The pre-school backgrounds of some children apparently do not prepare them adequately for the demands of formal schooling. Thus it would seem important to find suitable procedures for the early detection of such "high risk" children and make such procedures available to teachers of kindergarten and grade one children. In developing suitable programs, there appear to be three basic tasks facing the researcher, viz:

1. To determine those elements of a child's pre-school experience which are essential prerequisites for tasks which are part of a grade one program;
2. To develop devices which can meaningfully and economically assess such school relevant attributes;
3. To empirically determine which devices, singly and in combination, can most efficiently predict performance in various first grade tasks.

In addition, an examination of screening devices may be able to incorporate an assessment of factors purported to be characteristic of older children with learning problems, by means of tasks modified for administration to preschool children.

Having decided to use an intuitive approach to select an appropriate range of tasks, the test constructor may discover many devices which merit inclusion in an initial battery. In any event, the tasks of refinement and standardization of a screening battery will be likely to proceed along empirical lines rather than along intuitive or theoretical ones. As a test battery is progressively refined, its scope and limitations become more apparent. The test constructor may initially assume that his instrument is useful if it can provide reliable information not available from other sources. If the instrument can meet acceptable standards of reliability and validity after extensive field use, its value is greatly enhanced. Many psychological tests in current use have been developed along intuitive and empirical lines and have later provided useful data of heuristic value for theory development.

Although an intuitive/eclectic approach to test construction and an empirical approach to test development are defensible on the grounds of both precedence and expediency, it is important to note certain difficulties in these approaches. Reductionism is clearly involved in any attempt to explain a functioning "whole" in terms of the

interaction of its assumed parts. Secondly, the approaches suggested as useful in the present investigation are essentially correlational techniques. While inter-relatedness may be indicated in correlational analysis, it is not meaningful to infer a causal relationship between related variables.

Despite limitations, it would appear to be a useful step towards the effective screening of "high risk" children, to obtain reliability and validity data for potentially useful screening devices.

The subsequent review of literature addressed itself to an examination of screening devices and programs, and in addition sought to examine factors correlated with learning difficulties of older students, with a view to incorporating such information into an initial screening battery for preschool, kindergarten and beginning grade one children.

CHAPTER III

REVIEW OF LITERATURE AND DERIVATION OF HYPOTHESES

Several tests have been published recently which are alleged to assess factors related to performance in early school grades. To date, however, there has been very little published concerning the uses and limitations of many of these instruments.

An attempt has been made in this review to critically examine the literature pertaining to screening tests appropriate for grade one children. Additionally, material has been reviewed relevant to factors related to learning problems in older children, a source of screening tasks suggested by Bateman (1966).

For organizational purposes, it was expedient to arbitrarily separate literature relating to published screening or diagnostic devices from literature relating to correlates of learning problems.

Review of Published Screening Tests

First Grade Screening Test

The First Grade Screening Test (Pate & Webb, 1966) was designed to identify kindergarten or beginning grade one children who may be inadequately prepared for formal schooling. The FGST is alleged to assess emotional, intellectual and neurological functioning. Approximately 30 minutes is required for group administration involving the use of expendable answer booklets.

FGST norms were obtained from a stratified random sample of 5,534 first grade children. Eight months later, 770 of these children were tested on several achievement measures in order to obtain indications of the predictive validity of the FGST. Coefficients ranging from 0.60 to 0.79 were reported.

For prediction and screening purposes, various "cutting scores" were empirically determined. The manual includes descriptions of various types of communities which can be used as reference points for determining appropriate cutting scores. The manual reports test-retest coefficients of 0.84 and 0.82 for intervals of two and eight weeks respectively, while the inter and intra-scorer reliability coefficients are reported as 0.98. Apparently the FGST is undergoing nation-wide standardization at present.

Perceptual Forms Test

The Perceptual Forms Test (Sutphin, 1964) was developed as a project of the Winter Haven (Florida) Lions Club. As outlined in Sutphin (1964), the perceptual testing and training program is designed to contribute significantly to the development of reading and writing skills. The test involves copying seven geometric forms and may be administered either to individuals or to groups.

In her review of the Winter Haven program, Austen (1965) stated:

> There can be no question regarding the validity of measures of copying or reproduction ability as a predictor of school success (p. 849).

Austen's claim was based on studies by Lowder (1956) and Kagerer (1960). Although neither study convincingly demonstrated the value of the Winter Haven testing program as a predictor of school success, Austen's statement has been incorporated into advertising propaganda which is not supported by available evidence. It is somewhat surprising that Austen's cursory review was accepted by Buros (1965).

The Perceptual Forms Test (PFT) may be scored by either one of two methods superficially outlined in the test manual. Reliability and validity data are conspicuously absent from the manual.

Any attempt to effectively assess the value and limitations of the PFT must await further independent study in a variety of contexts.

The Marianne Frostig Developmental Test of Visual Perception

The Developmental Test of Visual Perception (Frostig, Maslow, Lefever & Whittlesey, 1964) evolved from the authors' attempts to evaluate and compare the visual perceptual skills of primary school children who were enrolled in California's "educational therapy" programs. Frostig's test consists of five subtests, assumed to effectively assess different aspects of visual perception, viz: eye-hand coordination; figure-ground discrimination; form constancy; perception of position in space; and spatial relations.

Frostig's test may be administered either to individuals or to groups and necessitates the use of expendable answer booklets by each examinee. By assuming that each sub-test measures a different factor in visual perception, the test constructors have developed a related program designed to provide specific individualized training in deficient areas.

The 1964 manual (Frostig, Maslow, Lefever & Whittlesey) has included norms for the 4.0 to 8.0 age range and has reported coefficients ranging from 0.40 to 0.50 between total test scores and measures of first grade reading achievement. Test-retest reliability coefficients ranging from 0.69 to 0.80 were reported for various restricted samples.

Allen, Jones & Haupt (1965) cautioned against using the battery for purposes other than those suggested by the authors. A factor analytic study by Corah & Powell (1963) failed to support the notion of sub-test independence, which as Mann & Phillips (1967) have noted, limits the value of programs designed to remediate specific visual perceptual problems indicated by sub-test performances.

Screening Tests for Identifying Children with Specific Language Disability

The Slingerland tests (1964) have been developed to provide teachers with instruments which can diagnose specific language problems in grades 1-4. As such, their purpose is more diagnostic than classificatory or predictive. Three levels are available

(grades 1-2; late 2-3; late 3-4), each having eight sub-tests.

Unfortunately details are not available in the manual concerning test construction, modification and standardization, nor is there any indication of reliability and validity data pertaining to previous administrations of the battery.

To date, little attention has been given to the Slingerland tests in psychological literature. Herman (1968) briefly outlined the development of Slingerland's summer training programs for teachers of dyslexic children. Paul (1967) suggested that the visual-memory-recall sub-test was moderately related to "reading." Paul's report is superficial, thus impeding any meaningful interpretation of the data offered in support of this claim.

Sub-tests in the Slingerland battery are labelled according to the test constructor's intuitive analysis of the skills which are assessed. It suffices to note that a child can fail a sub-test for reasons other than sub-average ability in the skills purported to be assessed. In the so-called "auditory" sub-tests, for example, failure may be due to poor visual discrimination, limited reading ability etc. Nevertheless, some of the sub-tests appear to be potentially valuable for screening if their interesting elements are retained yet modified so as to be appropriate for children who may not have attended school.

The Peabody Picture Vocabulary Test

The Peabody Picture Vocabulary Test (Dunn, 1965) is alleged to provide a measure of "hearing vocabulary by picture selection (p.32)." The test contains 150 plates each having four line-drawings. An examinee is asked to indicate which one of the four drawings best illustrates the stimulus, which is presented orally by the examiner.

Dunn claims that the test has several important advantages:

1. The test has high interest value and is therefore a good rapport establisher;
2. Extensive specialized preparation is not necessary for its administration;
3. It is quickly given in 10-15 minutes;
4. Scoring is completely objective and quickly accomplished in one or two minutes;
5. It is completely untimed and thus is a power rather than speed test;
6. No oral response is required;
7. Alternate forms of the test are provided to facilitate repeated measures; and
8. The test covers a wide age range (1965, p. 25).

Coefficients of stability ranging from 0.54 to 0.88, and coefficients of form equivalence from 0.61 to 0.97 were obtained in studies reported in the manual. Most validity studies were addressed to the question of what Dunn calls "congruent" validity (the comparability of PPVT with other vocabulary and intelligence tests). A few studies reported in the PPVT manual examined the relationships

between PPVT scores and scholastic achievement criteria (concurrent validity) while only two studies (Klaus & Starke, 1964; Moss, 1962) commented on the predictive validity of the PPVT.

Moss (1962) compared PPVT scores with scores obtained one year later on the Reading and Arithmetic subtests of the Metropolitan Achievement Tests (MAT), using a sample of 51 seven-year-old educable mentally retarded children (operationally defined as having Binet IQ scores of below 85). Correlations of 0.22 and 0.43 respectively were obtained. Klaus & Starke (1964) obtained correlations of 0.39, 0.35 and 0.39 between PPVT and MAT Word Knowledge, Word Discrimination and Combined Reading respectively, for a sample of 270 grade one children.

Dunn did not indicate the significance of the results in either study but suggested:

> Both of these studies were conducted on children at the beginning stages of reading and other subjects. Here visual discrimination and other factors are probably more important than hearing vocabulary in predicting school success. The PPVT should be a much better predictor from Grade 3 and on (p. 42).

One of the difficulties in using quotients at the age range of children beginning school is that no age discrimination finer than one year is possible for determining the IQ of children. All children between the ages of 5.6 and 6.5, for example, are grouped together. Thus, with the same raw score of say 50, there is a difference of 12 IQ points between CA's of 5.5 and 5.6 (Deal & Wood, 1968). The use of raw scores circumvents this difficulty to some extent.

Auditory Discrimination Test (Wepman, 1958)

Wepman's test is purported to identify children from five to eight years of age with auditory discrimination deficits. A subject is asked to indicate whether each of forty word pairs contains "same" or "different" words. Administration and scoring are straightforward and the test can be given in less than ten minutes per child. Di Carlo concluded his review in Buros (1965) by suggesting, on the basis of reviewed research:

> Auditory discrimination may represent only one of the contributory variables for performance in these (intelligence, hearing, speaking and reading) other areas... Nevertheless, for a quick, inexpensive, easy to score and accurate test of auditory discrimination, the Auditory Discrimination Test is highly recommended (p. 941).

The relevance of auditory factors for success in learning to read has been discussed extensively in the literature (Durrell & Murphy, 1953; Myklebust, 1954; Reichstein & Rosenstein, 1964; Sandstedt, 1964; Thompson, 1963; Wepman, 1960, 1964).

The Wechsler Intelligence Scale for Children: Digit Span Subtest

The Digit Span subtest of the WISC (Wechsler, 1949) has been suggested as a sensitive indicator of auditory perception/integration and distractibility difficulties (Krippner & Herald, 1963; Neville, 1960).

Although intended for administration as part of the complete WISC battery and therefore requiring some clinical expertise, the Digit Span sub-test by itself is objective, lucid and economical in

its administration and scoring. Accordingly, the Digit Span test appears appropriate to use as a teacher-administered screening instrument.

Review of Correlates of School Achievement

Bateman (1966) suggested that factors correlated with learning disorders seen in clinic populations could be potentially useful sources of tasks to include in a screening battery. It was considered appropriate to extend the present review to examine literature relevant to correlates of school achievement. In this way it was possible to review literature pertaining to factors such as language and teacher judgments and to include some multivariate studies undertaken with non-clinic samples.

Perceptual-Motor Factors

Kephart (1960, 1964) suggests that satisfactory performance on school-based tasks requires stability in a child's perceptual-motor world and that motor "patterns" must be established before uninterrupted interactions with the environment, and information gathering, are possible. The extensive and consistent interaction of the four basic patterns (viz: balance/maintenance of posture; locomotion; contact skills; and receipt and propulsion), enables stable processing of perceptual information. In Kephart's view, "ocular information is evaluated in terms of motor information (1964, p. 205)." Recent research (Rock & Harris, 1967) has suggested, however, that this may not

be the case. Vision, in fact, appears to train touch.

Laterality, Dominance, Body Image and Directionality Factors

A considerable body of literature pertaining to the relationships among factors such as dominance, laterality, directionality and body image has been reviewed by Zangwill (1962).

In recent years, strong claims have been made for renewed investigation of the interrelationships among these variables (Delacato, 1963; Phelps, 1965).

Phelps and Delacato have both suggested that the uniquely human activity of language is phylogenetically and ontogenetically preceded by hemispheric dominance, which is itself dependent upon consistent unilateral preference. Phelps suggests:

> Normally, dominance becomes set at about 18 months, which is of course the age at which the child begins to talk. Delayed setting of dominance thus results in speech delay (1965, p. 934).

In the proximity of the motor area of the brain responsible for control of hand and eye to the cortical areas involved in speech, Phelps sees a possible explanation for interdependence. He suggests that dominance can be successfully set by the use of restraining devices in 75-80% of children from two to four years of age, but that the percentage decreases with increasing age. Restraint seems unlikely to be successful with children over 12. The only comment offered in support of these claims is the vague statement ..."25 years of clinical experience (p. 941)." It has not been satisfactorily

demonstrated that cerebral dominance and unilateral preference are unvaryingly related (Luria, 1966) nor that either or both are necessary conditions for effective learning (Coleman and Deutch, 1964). Further, as Bateman suggests:

> Widespread educational application of findings in respect to cerebral dominance must await neurological clarification of many issues among them being (a) The relationship between cerebral dominance, which is a postulate, and laterality preference or observable behavior and (b) The feasibility as well as the desirability of intervention in the development of laterality (1966, p. 98).

Stephens, Cunningham & Stigler (1967) found no relationship between eye/hand preference patterns and reading achievement for a sample of 89 first grade children. Coleman & Deutch (1964), using matched groups of "normal" and "retarded" readers in upper elementary grades, noted no apparent relationship between laterality or right-left discrimination (as indicated by performances on the Harris Test of Lateral Dominance and the Benton Right-Left Discrimination Test respectively) and reading ability. Similar findings were reported by Balow & Balow (1964) with grade two children, and by Belmont & Birch (1965) with a random sample of 200 boys, 9-10 years of age, in Aberdeen, Scotland. This latter study noted, however, that confusion of right-left body schema was associated with lowest scores on tests of sequential reading.

Lovell, Shapton & Warren (1964), Rabinovitch & Ingram (1962), and Silver & Hagin (1960) have detected deficient right-left discrimination in some children with learning problems. Hendrickson & Muehl

(1962) have suggested the value of directional training for kindergarten children. Zaslow (1966) claimed to have successfully eliminated writing reversals by a training program which required each subject to cross the body mid-line and write on his opposite side. Collectively, these studies offer tentative evidence of a relationship between right-left confusion and learning problems.

Koos (1964) criticized studies which attempt to locate correlates of reading retardation in samples where 3 or 4 years of failure confounds any attempts to identify retrospectively those factors contributing to initial failure. Dominance anomalies, for example, may have existed at the time of initial confrontation with reading material yet may no longer be evident some years later. In a sample of 109 white, middle-class, primary grade children, Koos noted that reading difficulties were frequently associated with unestablished monocular preference. The presence of an intelligence threshold was suggested and some support was obtained for the hypothesis:

> Mixed dominant subjects of median or above median IQ (125 or greater) do not differ significantly in reading achievement from unilaterally dominant subjects in the primary grades, while mixed dominant subjects of below median IQ (less than 125) are signiticantly lower in reading achievement rank than unilaterally dominant subjects in these grades (1964, p. 156).

This finding can, at best, be considered as tentative. Luria suggests that ..."the degree of dominance varies considerably from subject to subject and from function to function ... both hemispheres

participate jointly in the performance of complex mental functions (1966, p. 87)."

The effectiveness of Wada's test (involving carotid artery injections of sodium amytal) as an indicator of hemispheric dominance has been demonstrated by Penfield & Roberts (1959) and Anzimirov & Karaseva (1967). The latter investigators and Luria (1966) noted that hemispheric dominance may sometimes differ for handedness and for speech.

Silver & Hagin (1960) noted a significant relationship between abnormal extension test performance (Appendix 1) and specific reading disability. For a sample of 150 reading disabled children (and 30 normal controls) between 8.6 and 14.0 years, defects in right-left discrimination were evident in 92% of the disability cases. Extension test abnormalities were apparent in 92% of the retarded readers but not in any of the control subjects. In another study (Silver & Hagin, 1967), 34 of 41 underachieving readers in a sample of 100 third and fourth graders, were noted to have extension test abnormalities. Of the 41 exhibiting abnormalities on the extension test, 39 were experiencing reading difficulties. In a 1964 study by the same authors, 24 of the subjects used in the 1960 investigation were retested. The extension test effectively differentiated the reading disability adults (median age 19 years) from the controls (median age 20 years).

These findings are made difficult to interpret by weaknesses evident in the available reports. In the 1960 study, for example,

subjects were referred initially for behavioral more than reading problems. Additionally, the criteria for delineating reading disability and the tests used as dependent variables are not reported.

Perhaps some pragmatic evidence of a possible relationship between mixed dominance and school learning problems can be inferred, although there appears to be an urgent need for some definitive research in this general area. For grade one screening purposes, it is important to examine the relationships between extension test and achievement test performances.

Maturational and Developmental Factors

"High risk" children have been described as lagging developmentally (Boder, 1966). It has been suggested that high risk children generally move and react in ways which are characteristic of chronologically younger children (Burks, 1960). To assist observers to determine general behavioral norms for various age levels, Ilg & Ames (1965) have constructed a developmental task battery based on years of experience with children at the Gesell Institute.

Sex-Linked Factors

Relevant literature consistently reports a high incidence of boys with learning problems (de Hirsch, Jansky & Langford, 1965; Ilg & Ames, 1965; Rabinovitch, 1964; Shimota, 1964).

Whether this phenomenon exists because of slower neurological maturation in boys, attitudinal factors, cultural mores, or sex-lined

factors operating singly or in combination, is open to conjecture.

Teacher Judgments

It is not surprising that several investigators have suggested that subjective ratings by experienced teachers are useful for identifying "high risk" children (Hall, 1966; Henig, 1949; Ilg & Ames, 1965; Kirk, 1966). Emotional and social immaturity are factors generally listed by teachers as important indicators of potential learning difficulties.

Pimm & McClure (1967) analyzed teachers' descriptions of some behavioral aberrations of emotionally disturbed children. A forced-choice, 100 item check list was developed and progressively refined until only 37 items were retained. A comparison of 31 emotionally disturbed elementary school children and 31 controls (matched on age, sex and IQ) revealed that 20 of the 37 items differentiated significantly between the two groups. It would be of interest to determine how effectively the check list, designed to identify children with potential behavior problems, can identify or predict "high risk" children. Bower & Lambert (1962) have developed an in-school screening procedure for identifying emotional handicaps.

Attwell, Orpet & Meyers (1967) examined the predictive value of teacher-ratings on ten observed behavioral categories, for a sample of 59 fifth grade children. "Attention" was the most successful predictor while "manual dexterity" and "effort displayed" were relatively

useful predictors. A nine-point (stanine) rating scale for each category was used.

Lee & Clark (1962) summarized some of the difficulties associated with teacher predictions:

> Unpublished studies by the authors (Lee & Clark) indicate that teachers differ markedly in their ability to estimate the reading readiness of their pupils. In 15 first grade classes the correlation coefficients of beginning grade one predictions with grade one achievement ranged from 0.16 to 0.72 (p. 6).

Additionally, it would seem likely that teacher ratings would improve as a function of the number of months spent with a given group of children, particularly with regard to grade one students for whom little background information is available.

Language Factors

In developing a pre-school program for "disadvantaged" children from lower socio-economic backgrounds, Bereiter & Engelmann (1966) have suggested the need to provide a verbal enrichment program to obviate what the authors claim to be a "cumulative deficit." The importance of language factors in school performance is forthrightly suggested:

> Verbal and reasoning abilities - which may be combined under the general rubric of ability to manipulate symbols - have been found to be the major factor in academic achievement throughout the school years (pp. 4-5).

The authors suggest that "the language deficiency of disadvantaged is ... the central handicap from which many of the others

(handicaps) derive (p. 13)." Several studies are cited in support of Bereiter & Engelmann's claim that, in essence, cultural deprivation is attributable to impoverished facility and experience with language.

The relationships between language and various indices of school achievement have been discussed by Loban (1963) and McCarthy (1964).

Multifactor Correlates of School Achievement

On the basis of their survey of the performances of over 1200 kindergarten children in Kansas, Haring & Ridgway (1967) suggested that a general language factor accounted for the greatest percentage of communality of variance, that physical performance and "intelligence" have little if any predictive value in identifying children with learning disabilities, and that the individual behavioral analyses done by teachers seemed as effective as group testing for screening purposes.

Barrett (1965) suggests that initial reading competence depends upon visual discrimination. In a stratified random sample of 724 grade one urban children, complete data were collected for 632 children, (331 boys and 301 girls). Independent variables were 7 visual discrimination tests, C.A., and intelligence (Lorge Thorndike). Dependent variables were two Gates Tests: Primary Word Recognition and Primary Paragraph Reading, with one year between pre- and post-testing. Multiple regression analysis indicated the best

predictors (in order) were Reading Letters, Pattern Copying, Word Matching and Reversals. Intelligence and C.A. were poor predictors. Six visual discrimination variables accounted for approximately 50% of the variance in the reading criteria. Obviously factors unaccounted for in this study share equal importance with visual discrimination skills in determining reading competence at the end of grade one.

De Hirsch, Jansky & Langford (1965) administered to 53 kindergarten-age subjects of normal intelligence, a battery of 37 perceptuomotor and linguistic tests. A profile on the tests, which were grouped according to processes assumed operative, was constructed for each child in terms of apparent strengths and weaknesses. At the end of grade two, a battery of achievement tests was administered and rank order correlation coefficients computed between "predictors" and "criteria." A Predictive Index was compiled consisting of the ten predictors which correlated significantly with grade two achievement. The tests on the Predictive Index were: Pencil Use; Bender-Gestalt Test; Wepman's Auditory Discrimination Test; Number of Words; Categories; Horst Reversals Test; and the Word Matching, Word Recognition I and II and Word Reproduction subtests of the Gates Reading Readiness Test.

The Index correctly predicted 10 of 11 "under-achievers" but also identified 4 false inclusions. If the purpose of the Index is screening for multidisciplinary testing and intervention, the

additional assistance for the false inclusions would be unlikely to be harmful.

The authors admit the preliminary nature of their findings. Details of a current study involving 400 children in New York City are not yet available. Some other tentative findings in their study were: that prediction was better for boys than for girls, that IQ was only a mildly useful predictor at the age of the subjects in the study, that modality preference and school achievement did not appear to be related and that prematurity and low birth weight appeared to be significant indicators of "high risk" children.

The likelihood of school difficulties among premature or low birth weight children has been discussed extensively by Pasamanick & Knobloch (1961) and Vernon (1967).

Derivation of Research Hypotheses

As previously indicated, the overall purpose of the present investigation was to examine the relationships between a variety of screening procedures and measures of grade one achievement. Throughout the study, four standardized achievement tests were used as dependent variables. The achievement tests selected (cf. Chapter IV) measured word recognition, comprehension of written instructions, spelling, and arithmetic concepts and skills.

From the preceding review of published tests and correlates of achievement, the following hypotheses were derived:

Hypothesis 1

There are significant, positive correlations between First Grade Screening Test performances and measures of grade one achievement.

This hypothesis was derived from claims made by the test constructors (Pate & Webb, 1966).

Hypothesis 2

There are significant, positive correlations between Perceptual Forms Test performances and measures of grade one reading achievement.

Sutphin (1964) and Austen (1965) have both suggested that PFT is a good predictor of reading in early school grades.

Hypothesis 3

a. There are significant, positive correlations between Marianne Frostig Developmental Test of Visual Perception total test scores and measures of grade one achievement.

b. There are positive, but non-significant correlations among subtests of the Marianne Frostig Developmental Test of Visual Perception.

Frostig, Maslow, Lefever & Whittlesey, (1964) have claimed that the battery is related to school achievement and that subtests measure relatively distinct aspects of visual perception.

Hypothesis 4

There are significant, positive correlations between total test scores on the Slingerland battery and measures of grade one achievement.

This hypothesis is based on claims by Slingerland (1964).

Hypothesis 5

There are significant, positive correlations between Peabody Picture Vocabulary Test raw scores and measures of grade one achievement.

Dunn (1965) has claimed that the PPVT is a moderately useful predictor of school success. Studies reviewed by Klaus & Starke (1962) and Moss (1962) tend to support Dunn's claim.

Hypothesis 6

There are significant, positive correlations between Auditory Discrimination Test performances and measures of grade one achievement.

This hypothesis has been inferred from observations in the literature concerning the relationship between auditory perception and reading (Durrell & Murphy, 1953; Myklebust, 1954; Reichstein & Rosenstein, 1964; Sandstedt, 1964; Thompson, 1963; Wepman, 1960, 1964).

Hypothesis 7

There are significant, positive correlations between scores on the Digit Span subtest of the Wechsler Intelligence Scale for Children and measures of grade one achievement.

Krippner & Herald (1963) and Neville (1960) observed relationships between WISC Digit Span and achievement.

Hypothesis 8

There are significant, positive correlations between a teacher's ordinal ranking of students and measures of grade one achievement.

This hypothesis was based on studies by Hall (1966), Henig (1949), Ilg & Ames (1965) and Pimm & McClure (1967) concerning the accuracy of teacher predictions.

Hypothesis 9

There are significant, positive correlations between extension test performance and measures of grade one reading.

In the face of the conflicting evidence reviewed concerning laterality and achievement, the hypothesis was derived from Silver & Hagin (1960, 1964, 1967).

Hypothesis 10

There are significant multiple correlations between optimal combinations of screening tests used in the present investigation and measures of grade one achievement.

After having administered a wide range of tasks all of which were allegedly related to school achievement, it seemed reasonable to assume that optimal combinations of screening tasks could be significantly related to school achievement, as suggested by the work of Barrett (1965), de Hirsch, Jansky & Langford (1965), and Haring & Ridgway (1967).

CHAPTER IV

RESEARCH DESIGN

The Sample

Data for the present investigation were obtained from first grade classes at three Edmonton schools, viz: Talmud Torah (Edmonton Hebrew School); Coronation Public School; and Glenora Public School. All students enrolled in grade one classes at each of these schools at the time of testing constituted the sample, there being two grade one classes at Talmud Torah and one grade one class at each of the other schools. A brief description of each school follows:

TABLE I

DESCRIPTION OF THE SAMPLE

School	Age Range[a]			Mean Age[a]			Number		
	Boys	Girls	Both	Boys	Girls	Both	Boys	Girls	Both
1. Talmud Torah	6.6-7.3	6.3-7.3	6.3-7.3	6.9	6.9	6.9	16	26	42
2. Coronation	6.4-7.5	6.4-7.2	6.4-7.5	6.10	6.9	6.10	14	13	27
3, Glenora	6.4-7.0	6.5-7.2	6.4-7.2	6.7	6.9	6.8	11	17	28

[a]Age in years and months

Talmud Torah offers programs for children at the nursery, kindergarten, primary and elementary levels. In each school day every child is taught on a half-day English, half-day Hebrew basis. From discussions with the principal and teachers, it appeared that most of the children came from financially secure families having high educational aspirations for their children.

Coronation and Glenora schools do not have kindergarten or preschool classes. Both schools are located in the same general geographical area of the city but students attending Talmud Torah travel daily from all parts of the city. From discussions with teachers and principals of both schools it appeared that Coronation school serves an area best described as middle-class-declining, while most children at Glenora came from home-owning families where fathers were employed in professional and skilled occupations.

It was not feasible to ascertain whether the combined sample could be considered representative of the Edmonton population of first grade students. From the preceding discussion, it is apparent that the sample was not wholly representative. Cross validation would be necessary before the results obtained in the present study could be considered applicable to grade one classes in other contexts.

The Test Instruments

For the initial phase of the study, a test was included in the battery of independent variables if it satisfied the following criteria:

1. The test appeared to assess functioning in an area related to school achievement;
2. The test could be unequivocably administered, scored and interpreted by grade one teachers;
3. Adequate performance by examinees did not depend upon previous experience with test-taking skills.

In their published forms, all eight sub-tests of the battery developed by Slingerland (1964) do not satisfy the third criterion. Nevertheless, the Screening Tests for Identifying Children with Specific Language Disability (Slingerland tests) are so little known that their inclusion in the initial phase of the study appeared justifiable.

Thirty-four independent variables and four dependent variables were investigated in the initial phase of the study (Table 2). However, as it was ultimately intended to develop a preschool screening procedure applicable for administration in many contexts, it was considered inadvisable to develop procedures for classifying children according to socio-economic criteria. While it is readily admitted that socio-economic status may be considered relevant to school performance, socio-economic considerations were not included in the investigation.

Subtests of the Marianne Frostig Developmental Test of Visual Perception (Frostig tests), the Wechsler Intelligence Scale for Children - Digit Span (WISC Digit Span) and the Slingerland tests were

TABLE 2

DESCRIPTION OF VARIABLES - PHASE 1 (TALMUD TORAH CHILDREN: N=42)

Variable	Reliability Coefficient	Hypothesized Factor
Independent Variables		
1. Age (in months)	.84[a]	
2. Sex	.86[a]	
3. Teacher Rating	.89[a]	
4. Hand-Eye Preference	.72[a]	laterality
5. Extension Test	.40[b]	hemispheric dominance
6. Right-Left Body Schema	.78[a]	awareness of body parts; left-right awareness
7. Writing Name	.84[a]	spelling; memory for letter and word forms
8. Auditory Discrimination Test	.91[c]	auditory perception
9. Number of Words	.61[a]	expressive vocabulary
10. PPVT (Raw Scores)	.54-.88[c]	recognition vocabulary
11. PPVT (IQ's)	.54-.88[c]	recognition vocabulary
12. WISC Digit Span (forwards)	.82[a]	freedom from distractibility and short term memory
13. WISC Digit Span (backwards)	.83[a]	freedom from distractibility and short term memory
14. WISC Digit Span (total)	.93[a]	freedom from distractibility and short term memory
15. First Grade Screening Test	.82-.84[c]	school readiness; various
16. Perceptual Forms Test	.74[a]	visual perception and drawing

17.-21.	Marianne Frostig Developmental Test of Visual Perception	.77[a]	1. eye-hand coordination
		.82[a]	2. figure=ground discrimination
		.88[a]	3. form constancy
	(5 sub-tests)	.68[a]	4. perception of position in space
		.81[a]	5. spatial relations
22.	Frostig battery (total test)	.69-.80[c]	visual perception
23.	Frostig battery (PQ's)	--	
24.-31	Screening Tests for Identifying Children with Specific Language Disability (Slingerland)	--	1. visual copying of words (far point)
		,85[a]	2. visual copying of words (near point)
		.90[a]	3. distractibility: memory for word pattern
	(8 sub-tests)	.84[a]	4. visual discrimination
		.89[a]	5. distractibility: recall of patterns
		--	6. auditory recall and letter association
			7. auditory perception and letter association
			8. auditory-visual association
32.	Slingerland battery (total test)	--	
33.	Home Language	--	
34.	Extension Test (Inter-rater check)	.40[b]	hemispheric dominance
	Dependent Variables		
35.	Word Recognition (WRAT)	.93[a]	oral reading: word recognition
36.	Following Directions (CAT-Form CC)	.85[a]	comprehension of meaningful prose
37.	Spelling (CAT-Form W)	.88[a]	spelling of words
38.	Arithmetic Concepts & Skills (MAT-Primary 1)	.83[a]	arithmetic concepts and skills

a least estimate of reliability from factor analysis communalities obtained from Talmud Torah data
b inter-rater reliability coefficient
c coefficients of stability reported in literature

examined as separate entities. Age, sex, hand-eye preference, right-left body orientation and home language were included as classificatory independent variables.

In addition to devices discussed in the preceding chapter, two additional measures, viz: Writing Name; and Number of Words (Appendix 1) suggested themselves to the writer following the analysis of pilot study results.

It was felt desirable to obtain measures of word recognition, comprehension of written material, spelling, and arithmetic concepts and skills, as four important areas of academic achievement in grade one. The following tests were chosen for this purpose:

1. Word Recognition

The Word Recognition subtest of the Wide Range Achievement Test (Jastak & Bijou, 1946) requires a child to read orally a list of letters then words, arranged in an ascending order of difficulty. Although admittedly part of an older test, this WRAT sub-test can discriminate well among students at all levels of word recognition ability. Unlike the equivalent Schonell test, the child who is unable to read words such as "do," "see," "cat," "milk" etc., may still attain a positive score by reading any of eight previous single letters.

2. Comprehension of Written Material

The Following Directions subtest of the California Achievement

Tests-Primary - Form CC (Tiegs & Clark, 1951) was used as a means of discriminating among first grade children with regard to following written directions. Although superceded, the Following Directions subtest was considered a useful device for complementing the reading skills assessed by the WRAT Word Recognition and for providing a relatively well graded series of tasks in less than one page of print.

3. Spelling

The California Achievement Tests - Primary - Form X Spelling subtest (Tiegs & Clark, 1963) provides a measure of spelling ability from a child's written attempts at a maximum of 20 words arranged in ascending order of difficulty.

4. Arithmetic Concepts and Skills

The Arithmetic Concepts and Skills subtest of the Primary 1 - Metropolitan Achievement Tests (Durost, Bixler, Hildreth, Lund & Wrightstone, 1963) was chosen largely because it assesses functioning on a wide range of grade one mathematical tasks, yet success on the test does not depend on reading ability.

Administrative Procedure

A validation study was conducted in March and April 1968, at school 1 (Talmud Torah) involving 42 grade one children (16 boys and 26 girls). Data were collected for all subjects on 34 independent variables and four criteria of academic achievement.

From scoring and analysis of children's efforts and classificatory data, it was apparent that several independent variables were inappropriate or irrelevant to the purposes of the study. Additionally, it seemed desirable to collect data from another sample of grade one children on a greatly reduced number of variables. In May 1968, the first grade children attending schools 2 and 3 were examined on 16 independent and four dependent variables. In this phase of testing, 55 children (25 boys and 30 girls) were involved.

As indicated, one of the purposes of the study was to examine how effectively a combination of screening devices could account for, or predict, variability in performance for each of the selected academic achievement criteria. For this type of question, multiple correlation analysis is appropriate if a large sample is used. Where a relatively small sample has been employed, the problem of shrinkage is significant (Nunnally, 1967). A multiple correlation coefficient (R) obtained from large numbers of independent variables and a small sample is substantially "biased upwards."

> When that ratio (number of subjects to the number of independent variables) is a 100:1, the bias is important... If there are as many as 10 variables from which the best several are to be selected, it will be wise to employ 500 or more persons in the study (pp. 164-165).

Administrative and cost factors made it impractical in the present investigation to obtain data from a sample even approaching 500. The combined sample of 97 first grade children from schools 1, 2 and 3 was admittedly a smaller than optimal sample for multiple

correlation analysis.

Eight independent variables were retained for multiple correlation computations (Table 3) following analysis of results from all three schools. Testing at school 1 did not include the drawing of a man as a separate test. However, as part of the First Grade Screening Test (FGST),all children had completed a drawing of a man. This FGST subtest was re-scored for analysis as a distinct variable. The children's efforts at writing their names were also re-scored prior to multiple correlation analysis (Appendix 1).

Statistical Procedure

All data for each phase of the present study were recorded in raw score form and transferred by the writer to IBM punch cards for computer analysis.

For the data from school 1, schools 1 and 2, and from the total sample, a separate correlation matrix was computed for boys, girls and combined groups, giving a total of nine correlation matrices to use in an attempt to examine hypotheses 1-9.

Estimates of reliability (for variables having unknown reliability) were obtained from factor analyses computed for each sample, by inspection of the unrotated factor matrix communalities.

For hypotheses 1-9, the choice of an appropriate alpha level for determining the significance of correlation coefficients, was influenced by Ferguson's comments (1966) concerning correlation

TABLE 3

ESTIMATES OF RELIABILITY: UNROTATED
FACTOR MATRIX COMMUNALITIES

Variable	School		
	1(N=42)	2 & 3(N=55)	1,2 & 3(N=97)
Independent Variable			
1. Age	.84	.68	.83
2. PPVT	.76	.78	.72
3. WISC Digit Span (F)	.82	.78	.61
4. Number of Words	.61	.58	.62
5. Wepman's A.D.T.	.80	.41	.62
6. Slingerland tests, #5	.89	.72	.60
7. Writing Name	.84	.75	.77
8. Drawing of a Man	--	.78	.84
Dependent Variable			
1. Word Recognition	.93	.85	.86
2. Following Directions	.85	.76	.73
3. Spelling	.88	.87	.80
4. Arithmetic	.83	.75	.67

coefficient attenuation due to moderately low reliabilities of relevant variables. Ferguson suggests that "errors in measurement tend to reduce the size of the correlation coefficient (p. 382)." It is evident (Tables 2 and 3) that the reported or obtained estimates of reliability are sufficiently low in some cases to cause attenuation of coefficients. Ferguson cautions:

> Low reliability may be compensated for by increase in sample size. An unreliable technique used with a small sample is, however, capable of detecting gross differences only, and the probability of not rejecting the null hypothesis when it is false may be high (1966, p. 385).

An alpha level of 0.05 was adopted. Where it was evident that a result was significant at beyond the 0.01 level ($p < .01$), this has been indicated. The significance of obtained correlation coefficients was examined using the appropriate "t" test, having N-2 degrees of freedom (Ibid, p. 187).

To examine hypothesis 10, multiple correlation coefficients were calculated for boys (N=41), girls (N=56) and combined sample data (N=97), according to the procedures discussed in Nunnally (1967, pp. 165-171). Prior to this, in an effort to minimize redundancy among independent variables, multiple correlation analyses were undertaken for both Talmud Torah and Coronation/Glenora data and least useful predictors omitted from multiple correlation analysis for the total sample. Each multiple correlation coefficient (R) was corrected for shrinkage (Nunnally, 1967, p. 164) and each obtained unbiased

estimate of the population multiple correlation ($\hat{R}$) was tested for significance using the appropriate "F" test with df_1=k (number of independent variables) and df_2=N-k-1 with an α = .05 (Ferguson, 1966, p. 401).

Decisions concerning significance of results were based on Talmud Torah data unless otherwise indicated. The Perceptual Forms Test, the extension test and some subtests of the Slingerland and Frostig batteries were not administered to schools 2 and 3, thus precluding significance testing of scores from all 97 children for hypotheses 2, 3, 4 and 9.

CHAPTER V

RESULTS AND DISCUSSION

Hypothesis 1

From test constructors' claims, significant and positive correlations between First Grade Screening Test and grade one achievement test performances were predicted.

It is apparent (Table 4) that FGST scores were not significantly related to reading and spelling measures but were moderately related to arithmetic scores. For schools 2 and 3, Following Directions and Spelling were moderately related to FGST scores ($p < .05$), accounting for approximately 12% of the variance.

When boys and girls' results were examined separately, the FGST was significantly related to all achievement criteria for girls at school 1 ($p < .01$), and to all except arithmetic for girls at schools 2 and 3 ($p < .05$). The FGST was not significantly related to any achievement measure for boys' scores. The range and types of items sampled by the FGST were apparently inadequate to discriminate effectively between higher and lower achieving grade one boys.

Some revisions of the FGST could be considered for future administrations. Vocabulary items, for example, require a child to encircle one out of eight illustrations on each page. It would seem that better use of booklets, an extended sample of vocabulary and a more reliable measure of vocabulary would result from presenting

three or four stimulus words per page, to be identified in turn by different symbols (eg. 0, -, X, +). A similar extension of items evaluating short-term recall may help discriminate more effectively among students. Valuable information is lost and discrimination is difficult for the item involving the drawing of a man. It may be useful to change the scoring criteria and award more points for this item.

TABLE 4

FIRST GRADE SCREENING TEST & GRADE ONE ACHIEVEMENT: CORRELATIONS

School	Group	N	Achievement Test				ra
			Word Recognition	Following Directions	Spelling	Arithmetic	
1	Boys	16	.03	.09	.14	.07	--
	Girls	26	.51**	.69**	.53**	.72**	--
	Both	42	.19	.30	.25	.34*	.83
2 and 3	Boys	25	.17	.20	.30	.25	.62
	Girls	30	.40*	.45*	.38*	.23	.74
	Both	55	.25	.35*	.35*	.24	.64

[a]Least estimate of reliability from unrotated factor matrix communalities.

*$p < .05$

**$p < .01$

As published, the FGST appears to be a moderately useful screening device for administration to grade one girls. Its value for kindergarten and preschool children was not examined.

Hypothesis 2

It was predicted that the Perceptual Forms Test and grade one achievement would be significantly related.

Results obtained (Table 5) failed to support this hypothesis.

TABLE 5

PERCEPTUAL FORMS TEST & GRADE ONE ACHIEVEMENT: CORRELATIONS

School	Group	N	Achievement Test				r^a
			Word Recognition	Following Directions	Spelling	Arithmetic	
1	Boys	16	.41	.45	.56*	.37	--
	Girls	26	.21	.17	.17	.20	--
	Both	42	.20	.27	.20	.29	.74

[a]Least estimate of reliability from unrotated factor matrix communalities.
*$p < .05$

If, as Austen (1965) suggests, copying tasks are useful for predicting school success, the Perceptual Forms Test could be expanded to discriminate more effectively between children at different levels of copying skill. The existing scoring criteria are poorly detailed.

A larger number of designs dichotomously scored (Appendix 3) could be researched.

In its present form, the Perceptual Forms Test appears to be of little value as a grade one screening test.

Hypotheses 3(a) and 3(b)

On the basis of the test constructors' claims, it was predicted that total test scores of the Marianne Frostig Developmental Test of Visual Perception would be significantly related to measures of grade one academic achievement and that non-significant correlations among subtests would be obtained.

The results of the present study (Table 6) offered strong support for hypothesis 3(a) for combined scores ($p < .01$), but not for hypothesis 3(b) which must be rejected.

Three additional observations from the results (Table 6) are of interest:

1. Subtest independence is more clearly evident from the boys' results;
2. The Frostig subtests appeared to be of more value as diagnostic entities for boys than for girls;
3. Frostig total test scores were significantly related to all measures of grade one achievement for boys ($p < .01$).

It would appear that the Frostig tests may be of some value for obtaining indications of general visual perception and line

TABLE 6

MARIANNE FROSTIG DEVELOPMENTAL TEST OF VISUAL PERCEPTION: CORRELATIONS: SCHOOL 1

Group	Item	Subtest					Total Test	Achievement Test				r^a
		1	2	3	4	5		W.R.	F.D.	Sp.	Arith.	
	1						48*	20	27	23	10	--
Boys	2	46					51*	49*	50*	46	17	--
	3	15	31				74**	59*	53*	58*	65**	
N=16	4	28	10	38			70**	59*	60*	63**	69**	--
	5	21	10	44			60*	26	18	35	22	--
	T	48*	51*	77**	70**	60*		74**	71**	76**	66**	--
	1						65**	39*	47*	40*	51**	--
Girls	2	35					46*	16	27	15	01	--
	3	21	05				67**	25	39*	20	39*	--
N=26	4	37	01	44*			78**	12	17	04	34	--
	5	66**	51**	33	45*		78**	43*	40*	38*	34	--
	T	65**	46*	67**	78**	78**		34	44*	28	46*	--
	1						54**	35*	39*	36*	19	77
Both	2	41**					51**	34*	39*	31*	05	82
	3	03	19				69**	42**	38**	35*	53**	88
N=42	4	24	01	37*			72**	21	27	13	43**	68
	5	37*	38*	36*	42**		72**	39*	33*	38*	28	81
	T	54**	51**	69**	72**	72**		49**	57**	42**	55**	

Note Decimals omitted

[a]Least estimate of reliability: Factor matrix communalities

*$p < .05$;

**$p < .01$.

drawing skills of grade one children. The pupils' answer booklets are expensive and the program involves considerable testing and scoring time. Remedial programs, based on processes assumed defective by low subtest scores, would appear to be of limited value, particularly for grade one girls.

Hypothesis 4

It was hypothesized that performances on Slingerland's Screening Tests for Identifying Children with Specific Language Disability would be significantly related to measures of grade one achievement.

Results obtained from the children at Talmud Torah school (Table 7) are highly significant and strongly support hypothesis 4.

Other observations are of interest:

1. All subtests were significantly related to achievement test scores ($p < .01$), all were significantly related to each other, having correlations ranging from 0.54 to 0.90 ($p < .01$), and all were significantly related to total test scores (correlations ranging from 0.78 to 0.95; $p < .01$);
2. Significant correlations were obtained between the Slingerland battery total scores and achievement measures for boys, girls and combined samples ($p < .01$).

It was suggested (cf. p. 34) that all eight subtests were inappropriate for preschool children. Subtests 2, 3, 4 and 5 appeared to be marginally appropriate from indications obtained from the

TABLE 7

SCREENING TESTS FOR IDENTIFYING CHILDREN WITH SPECIFIC LANGUAGE DISABILITY: CORRELATIONS: SCHOOL 1

Item	Subtest							Total Test	Achievement Test			
	1	2	3	4	5	6	7		Word Recognition	Following Directions	Spelling	Arithmetic
1								.78	.54	.52	.40	.36
2	.54							.79	.52	.58	.43	.53
3	.53	.55						.81	.58	.68	.47	.72
4	.59	.67	.76					.90	.62	.66	.52	.63
5	.64	.69	.75	.84				.93	.72	.76	.61	.65
6	.67	.78	.73	.86	.89			.95	.73	.82	.60	.70
7	.57	.64	.84	.88	.87	.87		.92	.71	.77	.62	.71
8	.68	.67	.77	.82	.83	.90	.87	.92	.66	.72	.48	.68

Note: All correlations significant ($p < .01$).

children in school 1. Subtests 2, 3, 4 and 5 were used at Glenora and Coronation schools (Table 8). As subtest 5 was significantly related to all measures of achievement ($p < .01$) and appeared to assess something akin to the test battery overall ($r = 0.93$), it was used for multiple correlation analysis.

Hypothesis 5

From indications in relevant literature, it was hypothesized that Peabody Picture Vocabulary Test raw scores would be significantly related to measures of grade one achievement.

From results obtained at school 1 (Table 9), only partial support for the hypothesis was indicated. Correlation coefficients were significant relative to Following Directions and Arithmetic only. For girls' results, however, PPVT was significantly related to all achievement measures, while correlations with all achievement test scores for boys were not significant. A similar trend was evident in the performances of children at schools 2 and 3. For the combined sample ($N = 97$), the hypothesis was supported. The PPVT was significantly related to all four achievement criteria ($p < .01$) and accounted for from 7-15% of the variance in achievement test scores.

The PPVT appears to be a moderately useful test for screening grade one children, particularly girls. The results from the present study seem to be consistent with Dunn's comments concerning the value of the test at this level (1965, p. 16).

TABLE 8

SCREENING TESTS FOR IDENTIFYING CHILDREN WITH SPECIFIC LANGUAGE DISABILITY: SUBTESTS 2, 3, 4, 5 & GRADE ONE ACHIEVEMENT: CORRELATIONS FOR SCHOOLS 2 AND 3

Subtest	Group	N	Achievement Test				r^a
			Word Recognition	Following Directions	Spelling	Arithmetic	
2	Boys	25	.32	.40*	.36	.42*	.75
	Girls	30	.36*	.19	.38*	.22	.81
	Both	55	.32*	.32*	.39**	.28*	.70
3	Boys	25	.35	.36	.37	.16	.52
	Girls	30	.40*	.40*	.52**	.20	.77
	Both	55	.37**	.37**	.43**	.16	.72
4	Boys	25	-.02	.20	.11	.32	.80
	Girls	30	.19	.20	.32	.12	.85
	Both	55	.07	.20	.23	.18	.76
5	Boys	25	.43*	.39*	.56**	.30	.80
	Girls	30	.55**	.60**	.56**	.59**	.80
	Both	55	.46**	.50**	.57**	.46**	.72
5	Total Sample	97	.54**	.59**	.56**	.53**	.60

[a]Least estimate of reliability from unrotated factor matrix communalities

*$p < .05$

**$p < .01$

TABLE 9

PEABODY PICTURE VOCABULARY TEST & GRADE
ONE ACHIEVEMENT: CORRELATIONS

School	Group	N	Achievement Test				r^a
			Word Recognition	Following Directions	Spelling	Arithmetic	
1	Boys	16	.35	.28	.33	.38	--
	Girls	26	.44*	.57**	.47*	.63**	--
	Both	42	.26	.35*	.28	.49**	.76
2 and 3	Boys	25	.41*	.21	.26	.23	.80
	Girls	30	.34*	.51**	.41*	.40*	.63
	Both	55	.39**	.30*	.28*	.30*	.78
1 2 and 3	Boys	41	.39*	.21	.27	.25	.72
	Girls	56	.39**	.53**	.42**	.50**	.56
	Both	97	.34**	.32**	.27**	.38**	.72

[a]Least estimate of reliability from unrotated factor matrix communalities

*$p < .05$

**$p < .01$

Hypothesis 6

On the basis of relevant literature, it was predicted that Wepman's Auditory Discrimination Test would be significantly related to grade one achievement test performances.

The results obtained (Table 10) offered only partial support for this hypothesis.

TABLE 10

AUDITORY DISCRIMINATION & GRADE ONE ACHIEVEMENT: CORRELATIONS

School	Group	N	Achievement Test				r^a
			Word Recognition	Following Directions	Spelling	Arithmetic	
1	Boys	16	-.05	-.02	-.19	.29	--
	Girls	26	.56**	.44*	.55**	.51**	--
	Total	42	.33*	.24	.33*	.41*	.80
2 and 3	Boys	25	.15	.27	.33	.24	.52
	Girls	30	.32	.59**	.33	.13	.76
	Total	55	.21	.45**	.32*	.17	.41
1, 2 and 3	Boys	41	-.04	.05	.12	.16	.83
	Girls	56	.32*	.45**	.35**	.18	.76
	Total	97	.14	.28**	.26**	.17	.62

[a]Least estimate of reliability from unrotated factor matrix communalities

*$p < .05$

**$p < .01$

The boys' results, the Auditory Discrimination Test was unrelated to all measures of grade one achievement, whereas for girls' results, significant correlations were obtained with reading and spelling measures. For the combined sample of 97 children, the Auditory Discrimination Test was significantly related ($p < .01$) to Following Directions and Spelling results only.

During classroom testing, the investigator noted that many children may have been handicapped inadvertently because of distance

or direction from the examiner. It is possible that a more accurate measure of auditory discrimination could be obtained by the use of a pre-recorded tape, leaving the examiner free to supervise group responses (as outlined in Appendix 4).

Hypothesis 7

It was hypothesized, on the basis of relevant literature, that WISC Digit Span total test scores would be significantly related to grade one achievement measures.

All achievement measures were significantly related to WISC Digit Span total test scores ($p < .01$), thus supporting the hypothesis.

The Digit Span test (Tables 11 and 12) appeared to be a useful screening test for boys and for girls. It is interesting to note that the Digits Forwards and Digits Backwards were correlated 0.54 ($p < .01$) and 0.27 ($p < .05$) respectively for results obtained from school 1 and schools 2 and 3 combined. Both parts were significantly related to total test scores ($r = 0.88$ and $r = 0.83$ for the separate samples). Either Digit Span Forwards or Digit Span Backwards would appear to be a useful indicator of Digit Span total test performance. Accordingly, in an effort to limit the number of independent variables included in the multiple correlation analysis, the Digit Span Forwards subtest was used. In retrospect, it would appear that Digit Span total test scores may have been a more thorough measure.

TABLE 11

WECHSLER INTELLIGENCE SCALE FOR CHILDREN - DIGIT SPAN & GRADE ONE ACHIEVEMENT: CORRELATIONS

School	Group	N	Achievement Test				r^a
			Word Recognition	Following Directions	Spelling	Arithmetic	
	Boys	16	.74**	.76**	.67**	.79**	
1	Girls	26	.39*	.48*	.43*	.46*	
	Both	42	.49**	.58**	.46**	.61**	.93
2	Boys	25	-.11	.63**	.52**	.57**	.96
and	Girls	30	.43*	.47**	.56**	.39*	.94
3	Both	55	.16	.54**	.55**	.49**	.95

[a]Least estimate of reliability from unrotated factor matrix communalities

*$p < .05$

**$p < .01$

During testing sessions, the investigator noted that discrimination among students may have been improved by scoring each attempt at every level of digit sequences. This procedure would enable children who can correctly repeat both sequences at each level to score more points than children who fail the first sequence at any level (cf. Appendix 4). It would be interesting to examine this possibility in a later study.

TABLE 12

WISC DIGIT SPAN SUB-TESTS & GRADE ONE ACHIEVEMENT: CORRELATIONS

School	N	Test	Item 1	2	3	4	5	6
1	42	1. D.S. (F)						
		2. D.S. (B)	.54**					
		3. D.S. (Total)	.88**	.88**				
		4. Word Recognition	.35*	.51**	.49**			
		5. Following Directions	.45**	.58**	.58**	.83**		
		6. Spelling	.29	.51**	.46**	.90**	.76**	
		7. Arithmetic	.63**	.43**	.61**	.59**	.74**	.49**
2 and 3	55	1. D.S. (F)						
		2. D.S. (B)	.27					
		3. D.S. (Total)	.83**	.76**				
		4. Word Recognition	.54**	.49**	.54**			
		5. Following Directions	.54**	.28*	.55**	.66**		
		6. Spelling	.53**	.36**	.49**	.85**	.76**	
		7. Arithmetic	.43**	.02	.40**	.46**	.56**	.51*

*$p < .05$

**$p < .01$

Hypothesis 8

On the basis of relevant literature it was hypothesized that there would be significant positive correlations between teachers' ordinal rankings of students and measures of grade one achievement.

In order to combine data from schools 2 and 3, it was

necessary to assume equality of rankings at both schools and consider each place in the ranking equal to its counterpart from the other school. Thus, the student ranked "1" at Glenora was considered to share position "1" on the ratings with the student ranked "1" at Coronation school. This spurious procedure resulted in a systematic depression of obtained coefficients, when compared with rank-order correlation coefficients, obtained for Glenora alone, (viz: 0.84; 0.79; 0.81; 0.78). The grade one teacher at Talmud Torah school ranked all 42 students together.

For all three schools (Table 13) strong support was obtained for the hypothesis. This is not surprising when one considers that the teachers concerned had been working with their groups for almost one school year. Although teachers' ratings were the best predictors of achievement test performance, they were not included in multiple correlation analysis.

TABLE 13

TEACHER RATINGS & GRADE ONE ACHIEVEMENT: CORRELATIONS

School	N	Achievement Test			
		Word Recognition	Following Directions	Spelling	Arithmetic
1	42	.90**	.80**	.82**	.59**
2 and 3	55	.68**	.71**	.77**	.61**

**$p < .01$

Hypothesis 9

From evidence reported by Silver & Hagin (1960, 1964, 1967) a significant relationship was hypothesized between extension test and achievement test performances.

This hypothesis was rejected on the basis of the results obtained (Table 14). Although the estimated coefficient of reliability was 0.89, the inter-rater (class teacher and experimenter) reliability coefficient was 0.40 for the combined sample of children from Talmud Torah school. Inter-rater reliability coefficients of 0.71 for boys and 0.17 for girls were obtained. The writer and the class teacher both observed that many students, particularly girls, did not consistently elevate one arm higher than the other. Two possible explanations for this observation are suggested;

1. Although the arm elevation may be related to hemispheric dominance, the evidence from the present study suggested that the extension test measured, at best, arm-muscle-tone and that the relevance of arm-muscle-tone to school achievement for this age group appeared to be negligible;
2. The tendency of 5-7 year old boys to engage in activities conductive to the development of better muscle-tone in one arm than in the other may account for the better inter-rater reliability obtained (0.71 as against 0.17 for girls). Experimenter's ratings of extension test performances correlated 0.52, 0.40, 0.61, 0.69 with achievement tests

used, while teacher's ratings correlated 0.32, 0.12, 0.49, -0.03. Girls' extension test performances correlated negligibly with achievement tests for both raters.

TABLE 14

EXTENSION TEST & GRADE ONE ACHIEVEMENT: CORRELATIONS FOR SCHOOL 1

Group	N	Achievement Test				r^a
		Word Recognition	Following Directions	Spelling	Arithmetic	
Boys	16	.52*	.40	.61	.30	
Girls	26	.31	.28	.21	.00	
Both	42	.29	.26	.21	.12	.89

[a]Least estimate of reliability from unrotated factor matrix communalities

*$p < .05$

The results obtained in the present study suggest that extension and achievement test performances are not related for the age group sampled. When poor inter-rater reliability was also considered, it appeared valueless to administer the extension test to children in the Coronation/Glenora sample. In retrospect, it appears regrettable that the additional observations were not made.

Hypothesis 10 (Multiple Correlation)

It was predicted that optimal combinations of variables, drawn from those hypothesized as relevant to schooling, would be significantly related to measures of grade one achievement.

The zero order correlation coefficients (Table 15) ranged from 0.10 to 0.60 for the total sample. Although age was not significantly related to reading or spelling, it was included because of its significant relationship with arithmetic performances.

Four variables in combination (writing of name, number of words, PPVT, Slingerland subtest 5) accounted for approximately 49% of the variance in Word Recognition scores. The $\hat{R}$ of 0.70 was significant ($p < .01$), thus supporting the hypothesis with regard to Word Recognition scores (Table 17). An $\hat{R}$ was calculated for boys and girls' results separately. Each $\hat{R}$ was significant ($\hat{R}$ boys = 0.66; $p < 01$ and $\hat{R}$ girls = 0.69; $p < 01$).

The significant predictors contain the elements of visual perceptual skill, familiarity with words and freedom from distractibility, which intuitively appear to be prerequisites for efficient reading at the grade one level.

Three variables (Table 17) in combination (Slingerland subtest 5, Digit Span Forwards, writing of name) were significantly related to performances in Following Directions ($\hat{R} = 0.67$; $p < 01$) accounting for approximately 45% of the variance, and thus providing support for the hypothesis. The combined correlation coefficient

for boys, after correction, was not significant ($\hat{R}$ = 0.55, $p > .05$), whereas the $\hat{R}$ for girls' results was highly significant ($\hat{R}$ = 0.80; $p < .01$). The Auditory Discrimination Test, the PPVT and the drawing of a man tests, all contributed to prediction for girls. The somewhat restricted nature of the samples used made it difficult to offer any clear explanation for this difference in prediction.

TABLE 15

PREDICTING GRADE ONE ACHIEVEMENT:
COMBINED SAMPLE: CORRELATIONS

Independent Variable	Achievement Test				r^a
	Word Recognition	Following Directions	Spelling	Arithmetic	
1. Age	.10	.10	.10	.30**	.83
2. PPVT	.34**	.32**	.27**	.38**	.72
3. D.S. (F)	.41**	.49**	.38**	.49**	.61
4. # Words	.51**	.42**	.46**	.24*	.62
5. Wepman	.14	.28**	.26**	.17	.82
6. Slinger-land 5	.54**	.59**	.56**	.53*	.69
7. Name	.60**	.55**	.51**	.47**	.77
8. Draw Man	.20*	.42**	.26**	.39**	.84

[a]Least estimate of reliability from unrotated factor matrix communalities

*$p < .05$

**$p < .01$

Four variables (Slingerland subtest 5, number of words, writing of name, Auditory Discrimination Test) were significantly related, in combination, to scores on the test of spelling ($\hat{R}$ = 0.64; $p < .01$), and accounted for approximately 41% of the variance (Table 18). Prediction was significant for boys' results ($\hat{R}$ = 0.66; $p < .01$) and for girls' results (R = 0.57; $p < .05$). Spelling and Word Recognition were significantly related (0.86) so that it is not surprising to note that a similar combination of predictors applied. All predictors except Age (Table 15) were significantly related to Spelling scores ($p < .01$).

From Table 18, one can readily observe that five variables (Slingerland subtest 5, Digit Span Forwards, PPVT, drawing of a man, age) were significantly related, in combination, to Arithmetic scores ($\hat{R}$ = 0.68; $p < .01$), accounting for approximately 46% of the variance. Prediction for girls' results was significant ($\hat{R}$ = 0.66; $p < .01$) but the $\hat{R}$ for boys' results was not significant (R = 0.58; $p > .05$).

It is interesting to note that prediction of Arithmetic performances was not significant for boys' scores. Teachers' ratings were somewhat less successful in predicting Arithmetic performance (Table 13). This may be accounted for partly by the fact that scores in arithmetic were only moderately related to other achievement measures (Table 16), suggesting that success in grade one arithmetic tasks may be relatively more difficult to predict than success in

reading and spelling. The battery used for multiple correlation analysis was moderately useful for predicting arithmetic performance for girls but was unsatisfactory for boys.

TABLE 16

GRADE ONE ACHIEVEMENT TESTS: CORRELATIONS FOR COMBINED SAMPLE

Achievement Test	Achievement Test			
	1	2	3	4
1. Word Recognition	1.00	.74	.86	.52
2. Following Directions	.74	1.00	.74	.63
3. Spelling	.86	.75	1.00	.48
4. Arithmetic	.52	.63	.48	1.00

Note: All correlation coefficients significant at .01 level.

One of the interesting features of the present study was the value of simple devices such as drawing a man, writing one's name, counting the number of words used in describing a picture, etc. Perhaps future studies of this nature can refine and cross validate these tests, in addition to including other simple tasks which may appear relevant.

TABLE 17

MULTIPLE CORRELATION: COMBINED SAMPLE: PREDICTION OF WORD RECOGNITION AND FOLLOWING DIRECTIONS

Test	Group	Independent Variable	Zero Order r	β Weight	Cumulative R^2 (%)	$\hat{R}$	F
Word Recognition	Boys N=41	# of Words	.63	.40	39.7		
		Writing Name	.57	.33	48.0		
		PPVT	.39	.25	53.5		
					R = .73	.66	3.1**
	Girls N=56	Writing Name	.63	.59	39.3		
		Wepman's test	.32	.29	50.1		
		# Words	.42	.28	57.6	.69	5.4**
					R = .75		
	Both N=97	Writing Name	.60	.38	35.8		
		# Words	.51	.28	47.3		
		PPVT	.34	.18	51.1		
		Slingerland #5	.54	.16	53.3		
					R = .73	.70	10.6**
Following Directions	Boys N=41	Slingerland #5	.55	.40	30.2		
		Digit Span (F)	.54	.39	43.1		
					R = .66	.55	1.7
	Girls N=56	Slingerland #5	.60	.22	36.2		
		Writing Name	.57	.42	47.7		
		Wepman's test	.45	.30	58.6		
		PPVT	.53	.26	64.0		
		Drawing of Man	.53	.24	67.7		
					R = .82	.80	10.4**
	Both N=97	Slingerland #5	.59	.36	34.7		
		Digit Span (F)	.49	.26	43.6		
		Writing Name	.55	.27	48.6		
					R = .70	.67	9.0**

*$p < .05$

**$p < .01$

TABLE 18

MULTIPLE CORRELATION: COMBINED SAMPLE:
PREDICTION OF SPELLING AND ARITHMETIC

Test	Group	Independent Variable	Zero Order r	β Weight	Cumulative R^2 (%)	$\hat{R}$	F
Spelling	Boys N=41	# Words	.66	.49	43.6		
		Slingerland #5	.58	.35	53.2		
					R = .73	.66	3.2**
	Girls N=56	Slingerland #5	.52	.27	27.0		
		Writing Name	.49	.38	35.6		
		Wepman's test	.35	.26	41.6		
					R = .65	.57	2.9*
	Both N=97	Slingerland #5	.56	.25	31.7		
		# Words	.46	.25	38.4		
		Writing Name	.51	.30	43.2		
		Wepman's test	.26	.17	45.7		
					R = .68	.64	7.6**
Arithmetic	Boys N=41	Digit Span (F)	.58	.44	34.0		
		Slingerland #5	.54	.37	45.8		
					R = .68	.58	2.1
	Girls N=56	Slingerland #5	.56	.30	31.7		
		Writing Name	.53	.33	41.9		
		PPVT	.50	.31	49.9		
					R = .71	.66	4.6**
	Both N=97	Slingerland #5	.53	.26	28.6		
		Digit Span (F)	.49	.28	39.1		
		PPVT	.38	.25	42.4		
		Drawing of Man	.39	.25	47.3		
		Age	.30	.16	49.9		
					R = .71	.68	9.4**

*$p < .05$

**$p < .01$

CHAPTER VI

SUMMARY OF RESULTS AND IMPLICATIONS FOR FURTHER RESEARCH

Summary of Results

1. The First Grade Screening Test appeared to be a moderately reliable and moderately useful measure of girls' grade one achievement. For boys' data, the FGST was unrelated to all four achievement criteria.

2. For all data from the Talmud Torah sample, the Perceptual Forms Test was unrelated to measures of grade one achievement.

3. The Marianne Frostig Developmental Test of Visual Perception was significantly related to grade one achievement for boys, girls and combined data in the Talmud Torah sample. Sub-test independence was noted for boys' data but not for girls' data nor combined data, where some significant relationships among sub-tests were observed. Sub-test independence and relationship to achievement were more clearly evident for boys' data.

4. The Slingerland test total scores were significantly related to grade one achievement measures for all data. All sub-tests were significantly related to achievement tests and to each other. In their existing form, the Slingerland tests are inappropriate for pre-school children because success in all sub-tests, particularly 1, 6, 7, 8, is dependent upon previous exposure to initial reading and writing skills.

5. The Peabody Picture Vocabulary Test appeared to have moderate reliability and was moderately related to measures of grade one achievement for girls and combined data. The PPVT appeared to be only slightly related to achievement criteria for boys.

6. For the combined grade one sample, Wepman's Auditory Discrimination Test was significantly related to girls' performances on grade one achievement tests. For boys' data, the Wepman test appeared to be unrelated to achievement criteria.

7. The WISC Digit Span total test scores were significantly related to measures of grade one achievement. The Digit Span Forwards subtest was significantly related to grade one achievement measures and to Digit Span total test scores.

8. Teachers' ratings were significantly related to all measures of grade one achievement. For all three samples, teacher ratings were the best predictors of achievement. Prediction of reading and spelling appeared to be slightly better than prediction of arithmetic performance.

9. For the Talmud Torah sample, the extension test proved to be unreliable and of little value as an indicator of school achievement.

10. Writing name, number of words, PPVT, and Slingerland sub-test 5, in combination, were significantly related to Word Recognition scores, accounting for approximately 49% of the variance. Age and Wepman's Auditory Discrimination Test were not significantly related to Word Recognition.

Slingerland sub-test 5, Digit Span Forwards and writing name were significantly related to Following Directions scores, accounting for approximately 45% of the variance. Prediction was not significant for boys ($\hat{R} = 0.55$) but highly significant for girls ($\hat{R} = 0.80$) where Wepman's Auditory Discrimination Test, PPVT and drawing a man test all contributed significantly to predictions.

Slingerland sub-test 5, number of words, writing name and Wepman's Auditory Discrimination Test, in combination, accounted for approximately 41% of the variance, and were significantly related to Spelling performance for boys, girls and combined samples.

Slingerland sub-test 5, Digit Span Forwards, PPVT, drawing a man, and age were significantly related to Arithmetic scores, accounting for approximately 46% of the variance. Prediction of boys' Arithmetic performances was not significant.

Implications for Further Research

The present study was handicapped, to some extent, by the limited scope of the testing sample upon which the results were based. Accordingly, it is suggested that a similar battery (modified as indicated in Appendix 4) be administered to large and representative samples of kindergarten, grade one and grade two students, in order to obtain additional information concerning uses and limitations of the battery and of specific subtests. Follow-up studies to obtain information concerning predictive validity are needed. Correlational

and factor analytic studies may provide insights into the independence and nature of subtests.

Individually tailored programs based on subtest strengths and weaknesses would be much more meaningful when more is known concerning what processes or skills are being tested by various subtests and how reliable these measures are. Intuitively, the modified battery discussed in Appendix 4 appears to assess a wide range of skills. More useful programs can be developed when the relationships among subtests for various age groups have been more clearly established and the administrative and scoring details have been examined for ambiguity and reliability.

Assuming that a modified battery of tasks can be developed and that it can assess an adequate range of skills relevant to success in academic subjects at the primary school level, some important practical implications follow. The battery could, for example, be used for the initial screening of school entrants. Those performing poorly in specific areas may require more specialist diagnosis. Subtest performances could be used as the bases for grouping of children for teaching or diagnostic purposes. The subtests may also prove to be of value as devices for evaluating the effectiveness of certain teaching programs or materials, and for evaluating progress in specific skills.

It is readily apparent that skills such as listening, writing, defining concepts, etc., continue to be important throughout schooling.

Tests such as the modified Auditory Discrimination Test, the PPVT, the drawing of a man etc., may be useful diagnostic devices for primary and elementary school children of all levels of ability.

An important feature of the present study was the highlighting of serious weaknesses in some currently published tests. It would seem most important that the uses and limitations of advertised tests be made known to school personnel. The present study, although limited in scope, raised important questions concerning the value of a descriptive review such as Austen's (1965) and concerning the ethics of publishing tests for educational purposes which have been inadequately researched and inadequately refined.

It was of interest to note the apparent usefulness of straightforward, teacher-administered devices. Undoubtedly other simple tasks can be considered which may provide useful insights for diagnosis or screening of young children.

Undoubtedly, the questions raised by the present study are of greater consequence than the results obtained.

REFERENCES

REFERENCES

Allen, R.M., Jones, R.W., & Haupt, T.D. Note of caution for the research use of the Frostig test with mentally retarded children. Perceptual and Motor Skills, 1965, 21, 237-238.

American Psychological Association, Council of Editors. Publication Manual of the American Psychological Association. (1967 Revision) Washington, D.C.: APA, 1967.

Anzimirov, V.L., & Karaseva, T.A. Hemispheric dominance and its determination. Voprosy Psikhologii, 1967, 13, 81-86. (Psychological Abstracts, 1967, 41, No. 9235).

Attwell, A.A., Orpet, R.E., & Meyers, E.C. Kindergarten behavior ratings as a predictor of academic achievement. Journal of School Psychology, 1967, 6, 43-46.

Austen, M. Perceptual forms test. In O.K. Buros (Ed.), Sixth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1965. P. 848.

Balow, I., & Balow, B. Lateral dominance and reading achievement in the second grade. American Educational Research Journal, 1964, 1, 139-143.

Barrett, T.C. Visual discrimination tasks as predictors of first grade reading achievement. Reading Teacher, 1965, 18, 276-282.

Bateman, B. Learning disorders. Review of Educational Research, 1966, 36, 93-119.

Belmont, L., & Birch, H. Lateral dominance, lateral awareness and reading ability. Child Development, 1965, 36, 57-71.

Bereiter, C., & Engelmann, S. Teaching disadvantaged children in the preschool. Englewood-Cliffs, N.J.: Prentice-Hall, 1966.

Bijou, S.W., Theory and research in mental (developmental) retardation. Psychological Record, 1963, 13, 95-110.

Boder, E. A neuropediatric approach to the diagnosis and management of school behavior and learning disorders. In J. Hellmuth (Ed.), Learning disorders. Vol. 2. Washington, D.C.: Special Child Publications, 1966. Pp. 15-41.

Bower, E.M., & Lambert, N. A process of in-school screening of children with emotional handicaps: A manual for school administrators and teachers. Princeton, N.J.: Educational Testing Service, 1962.

Bruner, J.S. Towards a theory of instruction. Cambridge, Mass.: Belknap Press, 1966.

Bryant, N.D. Group test of reading readiness. In O.K. Buros (Ed.), Fifth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1959. Pp. 676-677.

Bryant, N.D. Characteristics of dyslexia and their remedial implication. Exceptional Children, 1964, 31, 195-199.

Burks, H.F. The hyperkinetic child. Exceptional Children, 1960, 27, 18-26.

Cattell, P. The Detroit kindergarten test. In O.K. Buros (Ed.), Third mental measurements yearbook. New Brunswick: Rutgers University Press, 1949. P. 274.

Coleman, R.I., & Deutch, C. Lateral dominance and right-left discrimination: A comparison of normal and retarded readers. Perceptual and Motor Skills, 1964, 19, 43-50.

Corah, N.L., & Powell, B.J. A factor analysis of the Frostig developmental test of visual perception. Perceptual and Motor Skills, 1963, 16, 59-63.

Deal, T.N., & Wood, P.L. Testing the early educational and psychological development of children - ages 3-6. Review of Educational Research, 1968, 38, 12-18.

de Hirsch, K., Jansky, J.J., & Langford, W.S. The prediction of reading,spelling,and writing disabilities in children: A preliminary study. New York: Columbia University Press, 1965.

Delacato, C.H. The diagnosis and treatment of speech and reading problems. Springfield, Ill.: Thomas, 1963.

Di Carlo, L. Wepman auditory discrimination test. In O.K. Buros (Ed.), Sixth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1965. Pp. 940-941.

Dunn, L.M. Expanded manual for Peabody picture vocabulary test. Minneapolis, Minn.: American Guidance Service, 1965.

Durost, W.N., Bixler, H.H., Hildreth, G.H., Lund, K.W., & Wrightstone, J.W. Metropolitan Achievement Tests. New York: Harcourt, Brace & World, 1963.

Durrell, D., & Murphy, H. The auditory discrimination factor in reading readiness and reading disability. Education, 1953, 73, 556-560.

Ferguson, G.A. Statistical analysis in psychology and education. New York: McGraw-Hill, 1966.

Frostig, M., Maslow, P., Lefever, D.W., & Whittlesey, J.R. The Marianne Frostig developmental test of visual perception: 1963 standardization. Palo Alto, Calif.: Consulting Psychologists Press, 1964.

Gates, A.I. Gates reading readiness tests. New York: Teachers' College, Columbia University, 1942.

Guilford, J.P. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1965.

Hall, M.D. Family relationships of latency age boys with emotionally based learning inhibitions. In J. Hellmuth (Ed.), Learning Disorders. Vol. 2. Washington, D.C.: Special Child Publications, 1966. Pp. 175-191.

Haring, N.G., & Ridgway, R.W. Early identification of children with learning disabilities. Exceptional Children, 1967, 33, 387-395.

Hendrickson, L.N., & Muehl, S. The effect of attention and motor response pretaining on learning to discriminate B and D in kindergarten children. Journal of Educational Psychology, 1962, 53, 236-241.

Henig, M.S. Predictive value of a reading readiness test and teachers' forecasts. Elementary School Journal, 1949, 50, 41-46.

Herman, B. Teacher training as a multisensory technique for use with dyslexic children. Academic Therapy Quarterly, 1968, 3, 275-277.

Horst, P. Psychological measurement and prediction. Belmont, Calif.: Wadsworth, 1966.

Hunt, J.M. Intelligence and experience. New York: Ronald Press, 1961.

Ilg, F., & Ames, F.B. School readiness. New York: Harper Row, 1965.

Jastak, J., & Bijou, S. Wide range achievement test. Wilmington, Delaware: Charles L. Story, 1946.

Kagerer, R.H. The relationship of visual perceptual performance in early grades to reading level in grade four. Winter Haven, Fla.: Winter Haven Lions Publication Committee, 1960.

Kephart, N.C. The slow learner in the classroom. Columbus, Ohio: Merrill, 1960.

Kephart, N.C. Perceptual-motor aspects of learning disabilities. Exceptional Children, 1964, 31, 201-206.

Kirk, W. A tentative procedure for selecting bright and slow children in kindergarten. Exceptional Children, 1966, 33, 235-241.

Klaus, R., & Starke, C. Experimental revision of the Peabody picture vocabulary test as a predictor of first grade reading ability. Unpublished manuscript. In L.M. Dunn. Expanded manual for the Peabody picture vocabulary test. Minneapolis, Minn.: American Guidance Service, 1965, P. 37.

Koos, E.M. Manifestations of cerebral dominance in reading retardation in primary grade children. Journal of Genetic Psychology, 1964, 104, 155-165.

Krippner, S., & Herald, C. Reading disability among the academically talented. Gifted Child Quarterly, 1964, 8, 12-20.

Lee, M.J., & Clark, W.W. Lee-Clark reading readiness test. Monterey, Calif.: California Test Bureau, 1962.

Loban, W.D. The language of elementary school children. Champaign, Ill.: National Council of Teachers of English, 1963.

Lovell, K., Shapton, D., & Warren, N.S. A study of some cognitive and other disabilities in backward readers of average intelligence as assessed by a non-verbal test. British Journal of Educational Psychology, 1964, 34, 48-64.

Lowder, R.G. Perceptual ability and school achievement: An exploratory study. Unpublished doctoral dissertation. (Dissertation Abstracts, 1956, 16, No. 2205).

Luria, A.R. Higher cortical functions in man. New York: Basic Books, 1966.

Mann, L., & Phillips, W.A. Fractional practices in special education: A critique. Exceptional Children, 1967, 33, 311-317.

Maslow, P., Frostig, M., Lefever, D., & Whittlesey, J. The Marianne Frostig developmental test of visual perception (1963 standardization). Perceptual and Motor Skills, 1964, 19, 463-499.

McCarthy, J. The importance of linguistic ability in the mentally retarded. Mental Retardation, 1964, 2, 90-96.

Moss, J.W. An evaluation of the Peabody picture vocabulary test with the PMA and 1937 Stanford-Binet on trainable children. In L.M. Dunn. Expanded manual for the Peabody picture vocabulary test. Minneapolis, Minn.: American Guidance Service, 1965, P. 38.

Myklebust, H.R. Auditory disorders in children: A manual for differential diagnosis. New York: Grune and Stratton, 1954.

Neville, D. A comparison of the WISC patterns of male retarded and non-retarded readers. Journal of Educational Research, 1961, 54, 195-197.

Nunnally, J. Psychometric theory. New York: McGraw-Hill, 1967.

Pasamanick, B., & Knobloch, H. Epidemiological studies on the complications of pregnancy and the birth process. In G. Caplan (Ed.), Prevention of mental disorders in children. New York: Basic Books, 1961. Pp. 74-94.

Pate, J., & Webb, W. First grade screening test: Manual. Circle Pines, Minn.: American Guidance Service, 1966.

Penfield, W., & Roberts, L. Speech and brain mechanisms. Princeton, N.J.: Princeton University Press, 1959.

Phelps, W.M. The mixed dominant child. In C.H. Carter (Ed.), Medical aspects of mental retardation. Springfield, Ill.: Thomas, 1965. Pp. 930-941.

Piaget, J. The psychology of intelligence. New York: Harcourt Brace, 1950.

Pimm, J.B., & McClure, G. A screening device for the early detection of emotional disturbance in a public school setting. Exceptional Children, 1967, 33, 647-648.

Rabinovitch, R.D., & Ingram, W. Neuropsychiatric considerations in reading retardation. Reading Teacher, 1962, 15, 433-438.

Reichstein, J., & Rosenstein, J. Differential diagnosis of auditory difficulties: A review of the literature. Exceptional Children, 1964, 21, 72-81.

Rock, I., & Harris, C. Vision and touch. Scientific American, 1967, 217, 96-104.

Sandstedt, B. Relationship between memory span and intelligence of severely retarded readers. Reading Teacher, 1964, 17, 246-250.

Shimota, H.E. Reading skills in emotionally disturbed, institutionalized adolescents. Journal of Educational Research, 1964, 106-111.

Silver, A., & Hagin, R. Specific reading disability: Delineation of the syndrome and relationship to cerebral dominance. Comprehensive Psychiatry, 1960, 1, 126-134.

Silver, A., & Hagin, R. Specific reading disability: Follow-up studies. American Journal of Orthopsychiatry, 1964, 34, 95-102.

Silver, A., & Hagin, R. Specific reading disability: An approach to diagnosis and treatment. Journal of Special Education, 1967, 1, 109-118.

Slingerland, B.H. Screening tests for identifying children with specific language disability. Cambridge, Mass.: Educators' Publishing Service, 1964.

Stephens, W.E., Cunningham, E.S., & Stigler, B.J. Reading readiness and eye-hand preference patterns in first grade children. Exceptional Children, 1967, 33, 481-488.

Sutphin, F.D. A perceptual testing and training handbook for first grade teachers. Winter Haven, Fla.: Winter Haven Lions Research Foundation, 1964.

Thompson, B.B. A longitudinal study of auditory discrimination. Journal of Educational Research, 1963, 56, 376-378.

Tiegs, E.W., & Clark, W.W. California achievement tests: Complete battery. Los Angeles: California Test Bureau, 1951.

Vernon, M. Prematurity and deafness: The magnitude and nature of the problem among deaf children. Exceptional Children, 1967, 33, 289-298.

Wechsler, D. Wechsler intelligence scale for children. New York: Psychological Corporation, 1949.

Weener, P., Barritt, L.S., & Semmel, M.I. A critical evaluation of the ITPA. Exceptional Children, 1967, 33, 372-380.

Wepman, J.M. Auditory discrimination test. Chicago: Language Research Associates, 1958.

Wepman, J.M. Auditory discrimination, speech and reading. Elementary School Journal, 1960, 9, 325-333.

Wepman, J.M. The perceptual basis for learning. In H.A. Robinson (Ed.), Meeting individual differences in reading. Chicago: University of Chicago Press, 1964. Pp. 25-33.

Zangwill, O.L. Dyslexia in relation to cerebral dominance. In J. Money (Ed.), Reading disability: Progress and research needs in dyslexia. Baltimore: John Hopkins Press, 1962, Pp. 103-113.

Zaslow, R.W. Reversals in children as a function of midline body orientation. Journal of Educational Psychology, 1966, 57, 133-139.

APPENDICES

APPENDIX I

ADMINISTRATION AND SCORING CRITERIA FOR UNPUBLISHED SCREENING DEVICES

1. Teacher Rating

Class teachers were asked to list, in order of academic merit, all children in their classes without regard to age, sex or behavior. Judgments were to be based on subjective evaluations of current level of performance in academic subjects combined.

2. Hand-Eye Preference

Every child was asked to throw a ball at a target (both arms in turn), stir a glass of water, write name (both hands in turn), open a door, aim a toy pistol, squint through a peep-hole, and look through a paper cylinder.

3. Extension test

After demonstration, each child was asked to keep his/her arms straight, fingers spread and eyes shut, and to raise arms slowly from sides until parallel with each other and with the floor. For scoring purposes, no difference in elevation and elevation of the arm opposite to that used for writing were scored 0. Slight elevation of the writing arm was scored 1.

4. Right-Left Body Schema

Each child was asked to indicate his left hand, left knee,

right ear, left eye, then to indicate these body parts on the examiner. Following this, the examinees were required to touch right ear with left hand, put left hand on right foot and to touch right eye with left hand.

Scoring was dichotomized: No errors or one error scoring 1, while two or more errors were scored 0.

5. Writing Name

Children were asked to write their first and last names carefully and to check their efforts.

Scoring criteria were initially: 3 points for first name/3 for last name: - Total 6. 1 point off for incorrect spelling, 1 further point off for incomplete word or initial only, and all 3 points lost for no attempt.

This scoring system was modified as follows for inclusion in multiple correlation computations:

4 points maximum for first name,

4 points maximum for last name,

8 points maximum score.

First and last names scored separately.

Points were awarded as follows:

4 points if completely correct, 3 points if spelling was correct but an incorrect letter form or letter case was used, 2 points if an incorrect attempt at spelling was made, 1 point if an

undecipherable attempt or single letter only was made, 0 points if no attempt was made.

6. Number of Words

Children were shown a picture of a table setting and asked to help point to all of the things, colors, shapes and sizes they could see. A picture of a bear on a neutral background was presented and each child asked to use as many words as possible to describe everything he/she could see in the picture.

Raw scores were used for analysis and children were given 1 point for each descriptive word used. The test was untimed, thus introducing a situation in which persistence may have been relevant.

7. Drawing of a Man

Points were awarded for any indication of awareness of the following:

1. head; 2. body; 3. 2 arms; 4. 2 legs; 5. shoes or toes; 6. 2 eyes; 7. 2 ears; 8. hair or hair line; 9. neck; 10. mouth; 11. lips or teeth; 12. fingers, hand or fist; 13. nose; 14. eyebrows; 15. eyelashes; 16. eye detail; 17. elbow or bent arm; 18. waistline or belt; 19. clothing; 20. approximate head/body/arms/legs proportion.

APPENDIX 2

INDEPENDENT VARIABLES: CORONATION/GLENORA PHASE

1. Age in months
2. PPVT (raw score)
3. WISC Digit Span (Forwards)
4. WISC Digit Span (Backwards)
5. WISC Digit Span (Total)
6. Number of Words
7. Wepman's Auditory Discrimination Test
8. FGST
9. Slingerland tests: Sub-test 2
10. Slingerland tests: Sub-test 3
11. Slingerland tests: Sub-test 4
12. Slingerland tests: Sub-test 5
13. Frostig's Developmental Test of Visual Perception: Sub-test 3
14. Frostig test: Sub-test 4
15. Writing Name
16. Drawing a Man

NOTE: Variable 16 was not part of the original program of 34 independent variables for the Talmud Torah phase of the study although drawings were collected from Talmud Torah children and thus could be re-scored as a separate variable.

APPENDIX 3

DESIGN COPYING TEST

Stimulus Cards:

In clear black lines on 5" x 8" white blank cards. Administer as timed or untimed test, with or without distracting stimulus - (Appendix 4, #4).

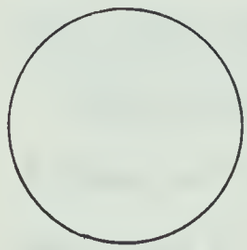
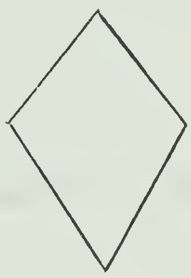
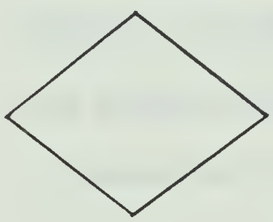
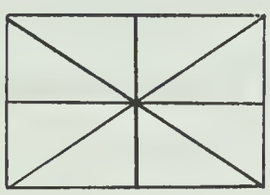
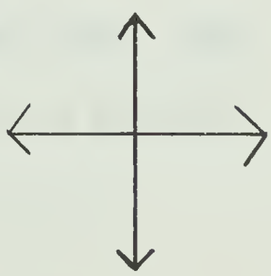
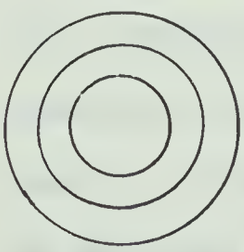
Answer Sheet:

Quarto size, ruled into four rows and three columns.

Scoring:

1 or 0. One point awarded if all elements are present, the child's effort resembles the stimulus more closely than any related figure and contains no additional or grossly distorted elements.

Stimulus Designs:

1. 	2. 	3. 
4. 	5. 	6. 
7. 	8. 	9. 
10. 	11. 	12. 

APPENDIX 4

SUGGESTED DEVICES FOR SCREENING KINDERGARTEN AND GRADE ONE CHILDREN

1. WISC Digit Span

Administration: Individual; 3-5 Minutes.

Dictate and score both stimulus digit sequences at each level. Award one point for each correct response. Discontinue after three consecutive failures.

2. Drawing of a Man

Administration: Group; 5-10 Minutes.

Score maximum of 20 points - as described in Appendix 1, #7.

3. Writing of Name

Administration: Group; 1-2 Minutes.

Child to write his/her first name and last name twice. Mark the better effort. Score maximum of 8 points - as discussed in Appendix 1, #5.

4. Design Copying Test

Administration: Group; 5-15 Minutes.

(a) Present each stimulus for 10 seconds while children have pencils on floor. Remove stimulus, ask children to pick up pencils and copy what they saw. Score 1 or 0 for each item - as described in Appendix 3.

(b) As an untimed test of copying skills, display each stimulus card for approximately two minutes. Score 1 or 0, maximum score of 12 points - as discussed in Appendix 3.

5. Wepman's Auditory Discrimination Test

Administration: Group; 5-15 Minutes.

Pre-record on quality tape recorder each pair of words (clearly pronounced). Present stimulus word pairs once only with 30 second pause between each pair. Children to indicate "same" or "different" by marking ✓ or X in each "box" of a stencil, ruled into 10 rows and four columns. Stop after each set of four pairs and check understanding of instructions and placement of answers, by quick survey of papers held under chins. Score maximum of 40, counting every item.

6. Peabody Picture Vocabulary Test

Administration: Individual; 5-15 Minutes.

Administer and score as prescribed in testing manual.

7. Number of Words

Administration: Individual; 2-5 Minutes.

Present colored photograph or illustration, administration and scoring as described in Appendix 1 #6.

8. Slingerland's Screening Tests for Identifying Children with Specific Language Disability -

(a) Subtest #3

Administration: Group; 5-10 Minutes.

(Marginally appropriate for kindergarten children)

Administration and scoring as described in testing manual.

(b) Subtest #5

Administration: Group; 5-10 Minutes.

(Not appropriate for preschool or kindergarten children)

Administration and scoring as described in testing manual.